Zadeh L.A., Tufiş D., Filip F.G., Dziţac I.(Eds.),
From Natural Language to Soft Computing: New Paradigms in Artificial Intelligence,
Editing House of Romanian Academy, Bucharest, ISBN 978-973-27-1678-6, 2008, pp. 84-99

Neuro-Fuzzy Modeling of Event Prediction using Time Intervals as Independent Variables

Simona Dziţac, Ioan Felea, Ioan Dziţac and Tiberiu Vesselenyi

Abstract: Estimation of possibility of future events is one of the main issues of predictive maintenance in technical systems. The paper presents an application of neuro-fuzzy modeling in reliability analysis. For this application we first consider some aspects of neuro-fuzzy modeling using the MATLAB programming environment and the mathematical model of event prediction. Than we succeed with prediction of time intervals at which events can occur using the registered data from a group of power energy distribution center maintenance databases. Conclusions are drawn regarding method applicability and future improvements.

Keywords: neuro-fuzzy modeling, prediction, membership function.

1 General considerations

Prediction of parameter values related to halts, events or failures of power distribution nets is one of the major simulation methods in the field of reliability. In this paper, the authors had searched o method which can generate a prediction model based on measurements of parameter values of the system. Such a model can be developed using neuro-fuzzy simulation methods [2, 3, 4, 5, 7].

Fuzzy reasoning has the ability to deal with uncertain information, while neural nets can be developed based on input-output data sets. Neuro-fuzzy algorithms combines the benefits of both methods.

The goal of this article is to present how neuro-fuzzy models can be used in the prediction of some events in the area of power distribution systems based on time intervals between two events. The time interval analysis can be then used to study how often events will occur in the future. The case study had been developed using the “events” tables from “Electric Energy Quality Indicators” database [1, 6, 8]. Prediction of time intervals at which events can occur had been made for the Center of Electric Energy Distribution of the City of Oradea, using the MATLAB environment as software development tool.

2 Neuro-fuzzy modeling

The base idea of neuro-fuzzy modeling is that the fuzzy system parameters (inference rule set, fuzzy membership functions) are established by learning (training) known input-output data sets. The learning methods are similar with those used in the case of neural nets.

The way in which this method is used is similar with the way in which the majority of identification techniques are used:

- in the first step a hypothesis is generated on the way that the system should operate and an initial model is defined (by associating membership functions and a set of rules).

- in the second step, input-output data sets are selected for training.
- in the third step using the selected data the adjustment of fuzzy membership functions is made using the neuro-fuzzy and accounting for corresponding minimal error criterion.
- in the final step, a testing data set is chosen and the developed system is tested with this set.

If the obtained results for the testing data set is satisfying then a series of methods exist in order to optimize the model:

a. increasing the amount of data in he training set;

b. increasing the training epochs number;

c. decreasing the increment of membership function adjustment algorithm.

In general, this type of model is working well if the training data can reflect representative characteristics of the modeled system. In practice the measured data set are not always reflecting the whole number of situations, which can occur and a certain value of uncertainty must be accepted.

System testing is made using a set of data different from the training data and the mean squared error obtained during the training epochs and the testing algorithm. If this error converges than the training is supposed to be correct.

Computing of adjustment parameters of membership functions is made with the help of a gradient vector, which optimizes the correlation between the neuro-fuzzy model and the real system specified by the input-output data. Once the gradient vector found, an optimization algorithm defined by minimizing the mean square error of real and modeled data is applied. The membership function parameter adjustment then is made using a hybrid algorithm based on the least square method and error back-propagation [2, 4, 5, 7].

In order to develop neuro-fuzzy models in MATLAB environment, the ANFIS module (Adaptive Neuro-Fuzzy Inference System) is used. ANFIS generates a self-adjustable fuzzy inference system by training with input-output data sets. The ANFIS module is called from MATLAB programs by the “anfis” command, which bears the following set of parameters [3]:

- number of training epochs;
- training error limit;
- initial value of optimization step;
- optimization step decrease rate;
- optimization step increase rate;

The training process stops when the number of prescribed epochs or the training error limit is reached.

The “anfis” command returns an informational structure, which is the neuro-fuzzy model of the process as it was given by the training data set.

3 Event prediction using neuro-fuzzy models

In the case that the goal is the prediction of future values based on a set of measured values reflecting a certain anterior evolution of a system, the use of a predictive model based on input-output data sets is somehow unusual, because only one set of data is at disposal.

That is the case of the values contained in time series, covering a domain of $[0, t]$, which values are to be used to predict the systems behavior at time $t + P$.

A convenient method for such prediction is to consider a series of N values from the known interval with the step $\triangle$:

$$(x(t-(N-1)\triangle), \ldots, x(t-\triangle), x(t)) \tag{1}$$

If, for example $\triangle = P$, is chosen, for each value of the known series, at time t the $w(t)$ vector can be defined as:

$$w(t) = [x(t-\triangle)x(t-\triangle)x(t-\triangle)x(t)] \tag{2}$$

The $w(t)$ vector can be used as a set of training input data for the neuro-fuzzy model. The output data set will be predicted (or estimated) as $s(t)$:

$$s(t) = x(t+\triangle) \tag{3}$$

As usual, the data set is divided in two intervals: the first interval is the training data and the second interval will be used as testing data.

In the training data, for each value the $w(t)$ value will be computed and the $w(t)$ vector will be considered as the input set and the second interval will be considered as the testing data set.

4 Application regarding time interval prediction at which events can occur in an electric power supply system

In the "Events" table of "Electric Energy Quality Indicators" database [1, 6, 8] there are recorded the dates of events at every consumer.

In the neuro-fuzzy predictive method, presented in the preceding paragraph, the algorithm had generated 16 rules with 104 parameters of the membership functions (considering "Gaussian" type membership functions). In order to reach satisfactory results is important that the number of training data be twice larger than the number of parameters of the membership functions. The case of the database used, a minimal set of 200 training data could be used, only in the case of "Electrica Nord Transilvania Oradea" company for a time interval of 6 years. Considering a predictive analysis on subdivisions (centers) of three distinct studies had been made for: Oradea, Stei and Alesd cities. From the reliability point of view it had been considered efficient to analyze the time between two events. So future consecutive event occurrence intervals can be estimated (or how frequently such events will occur in the future).

In order to fulfill the prediction task under the MATLAB environment, database information (initially worked out in EXCEL), had been decrypted with a special module [1]. Then the main interest fields were sorted (primary index is the date and time of the event and secondary index is the center name [1]).

From the sorted data, the number of days between two incidents had been counted and the neuro-fuzzy prediction was made for every center with the help of the software realized by the authors [1].

The parameters of the ANFIS command were as follows:

- number of training epochs = 150;
- training error limit = 0;
- initial value of the optimization step = 0,0001;
- the decrease rate of the optimization step = 1,9;
- the increase rate of the optimization step = 2,1;

For Alesd center the training data set and the testing data set was 200 values each and for Stei and Oradea centers was 300 values each.

Results of program running are presented in diagrams of figures 1-18. If the event occurrence diagram is required then it can be extracted from the predicted event period data using a specially designed program module, which can rebuild the time scale and represent the eventless periods as zeros and the event as values of ones. Using this function the diagrams in figures 19-21 were obtained.

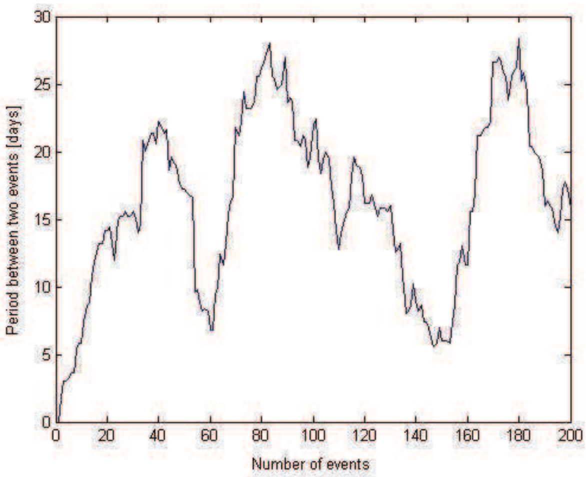

Figure 1: Training data for center Aleşd.

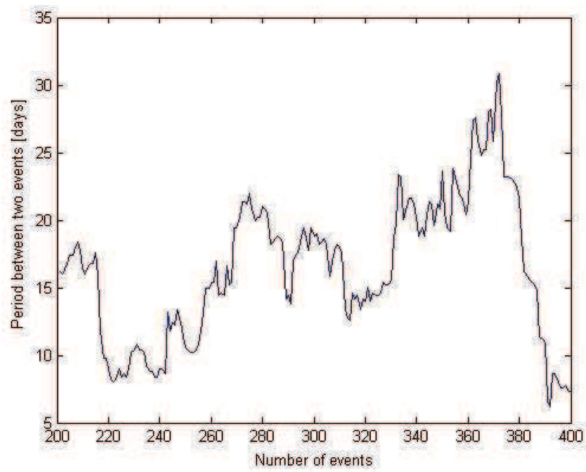

Figure 2: Testing data for center Aleşd.

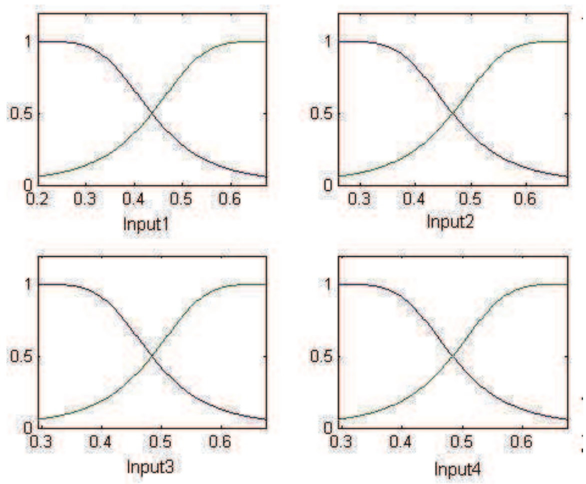

Figure 3: Initial membership functions for centre Aleşd.

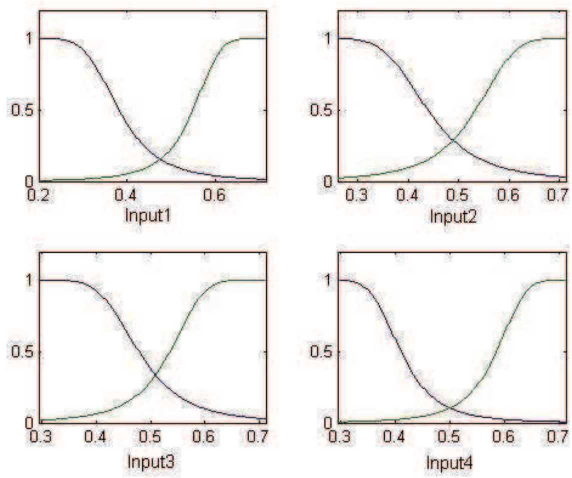

Figure 4: Adjusted membership functions for center Aleşd.

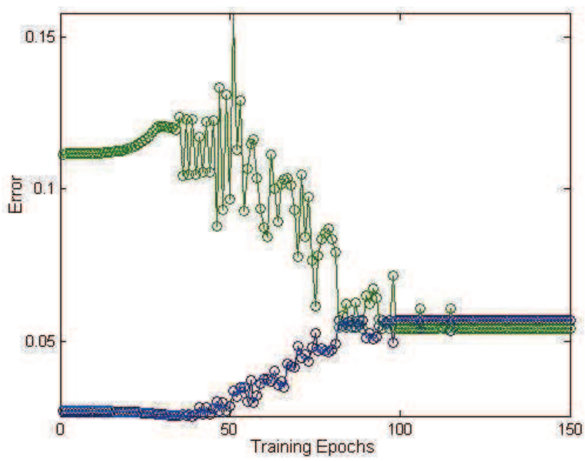

Figure 5: Comparison of training (blue) and testing (green) errors for center Aleşd.

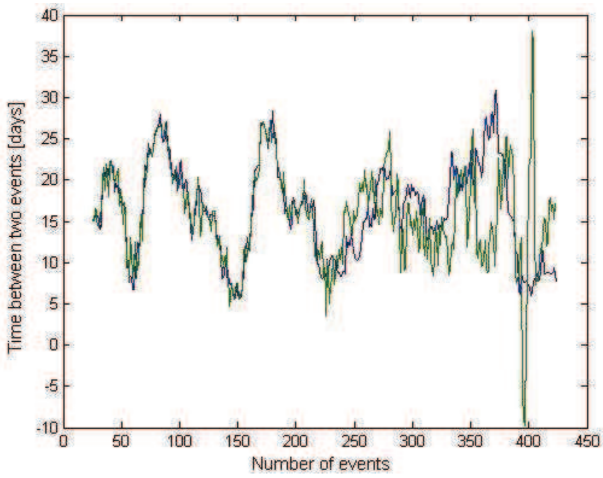

Figure 6: Event period prediction diagram real values (blue); simulated values (green) center Aleşd.

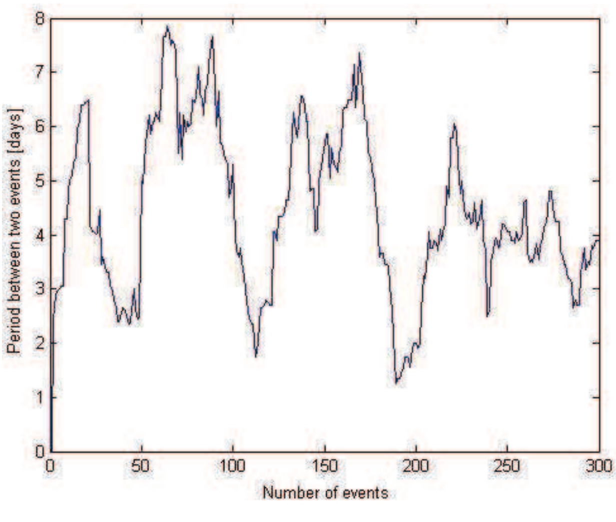

Figure 7: Training data for center Stei.

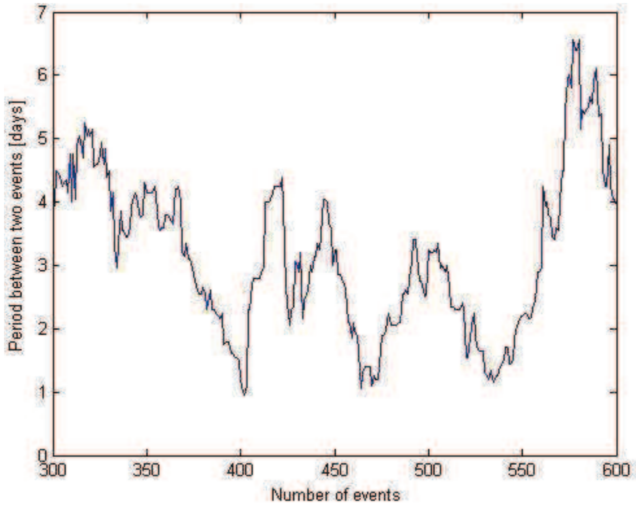

Figure 8: Testing data for center Stei

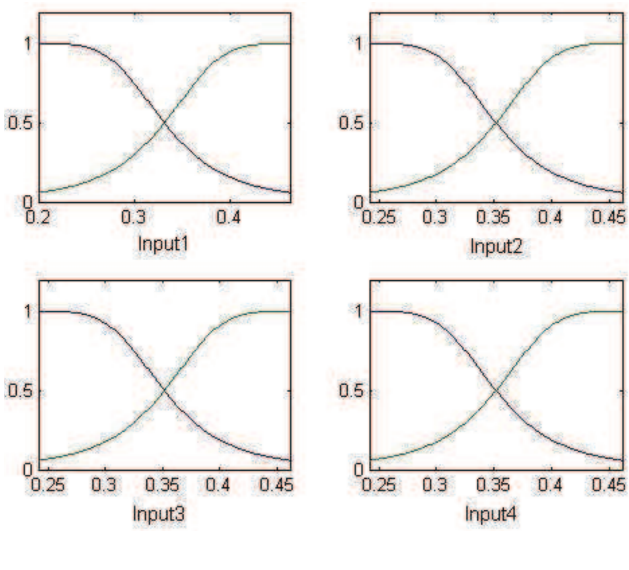

Figure 9: Initial membership functions for center Stei.

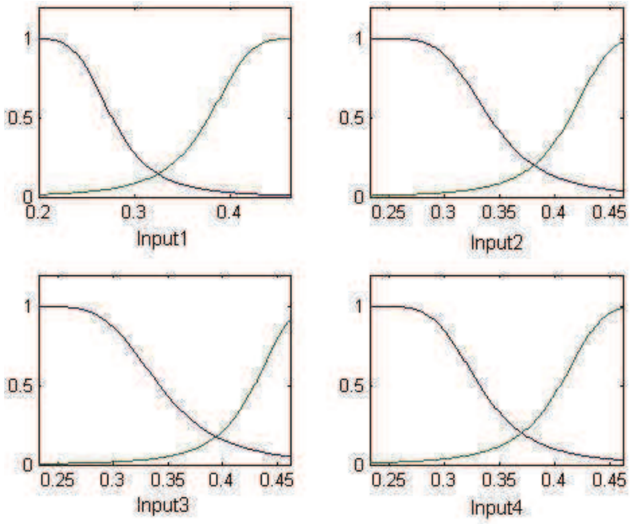

Figure 10: Adjusted membership functions for center Stei.

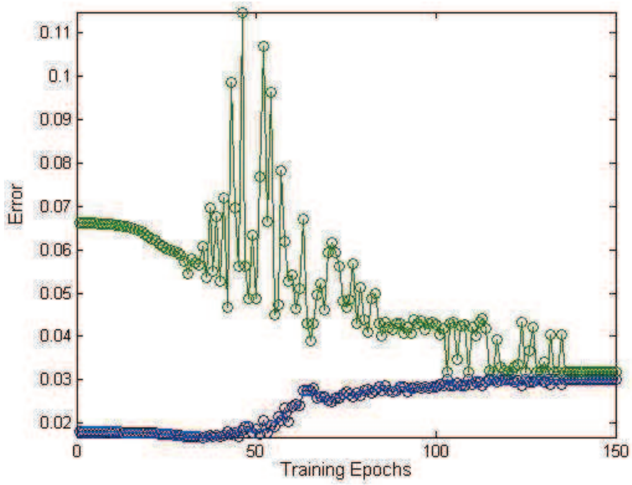

Figure 11: Comparison of training (blue) and testing (green) errors for center Stei.

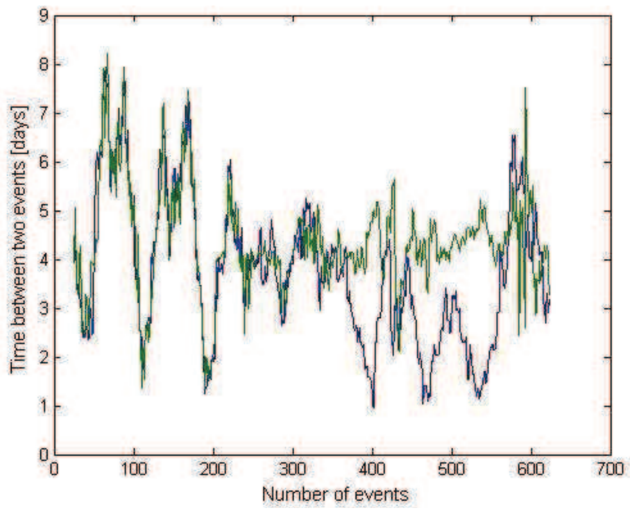

Figure 12: Event period prediction diagram real values (blue); simulated values (green) center Stei.

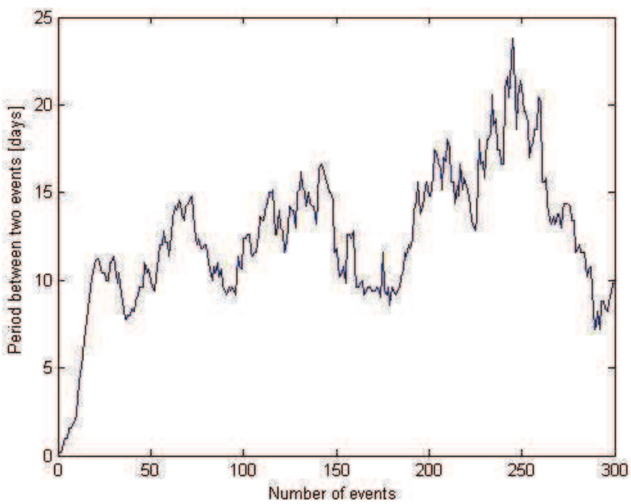

Figure 13: Training data for centre Oradea

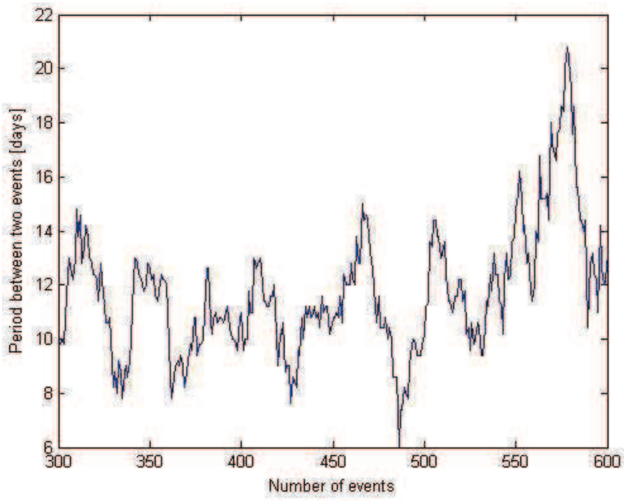

Figure 14: Testing data for centre Oradea

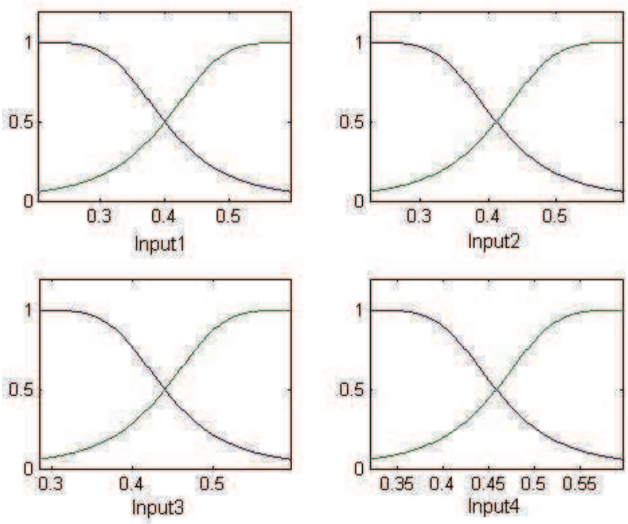

Figure 15: Initial membership functions for centre Oradea

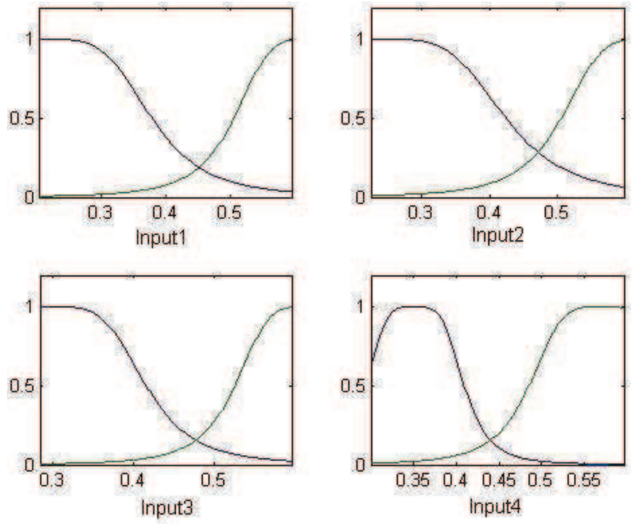

Figure 16: Adjusted membership functions for centre Oradea

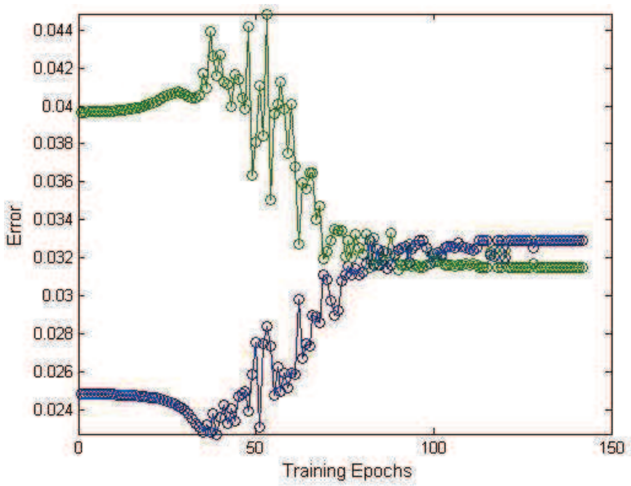

Figure 17: Comparison of training (blue) and testing (green) errors for centre Oradea

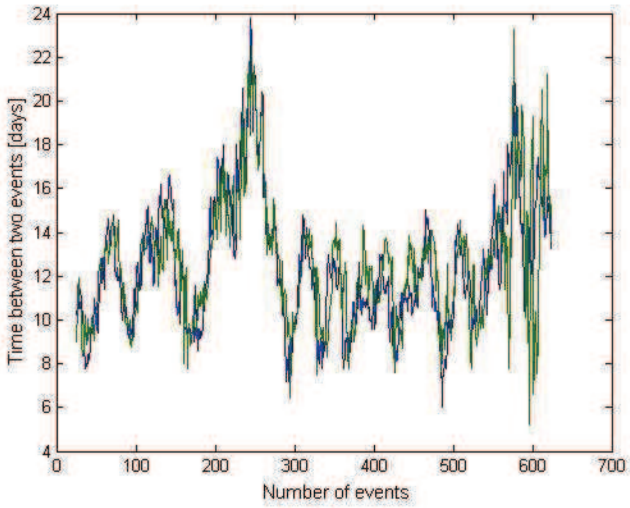

Figure 18: Event period prediction diagram real values (blue); simulated values (green) center Oradea.

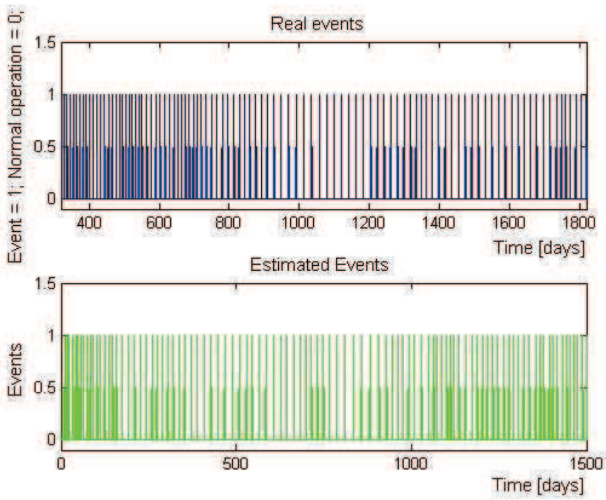

Figure 19: Event occurrence diagram for centre Aleşd

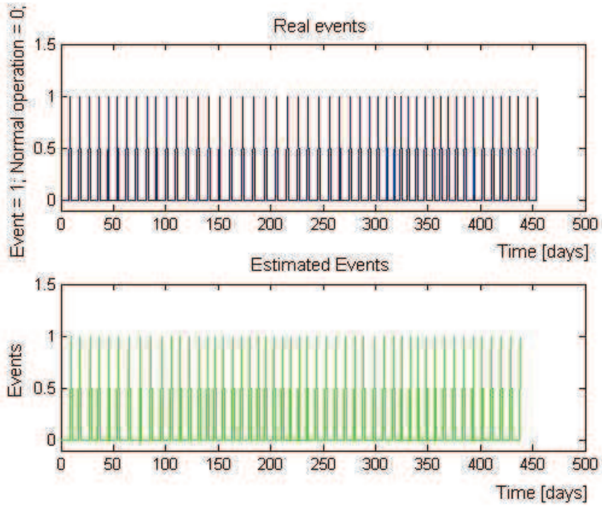

Figure 20: Event occurrence diagram for centre Ştei

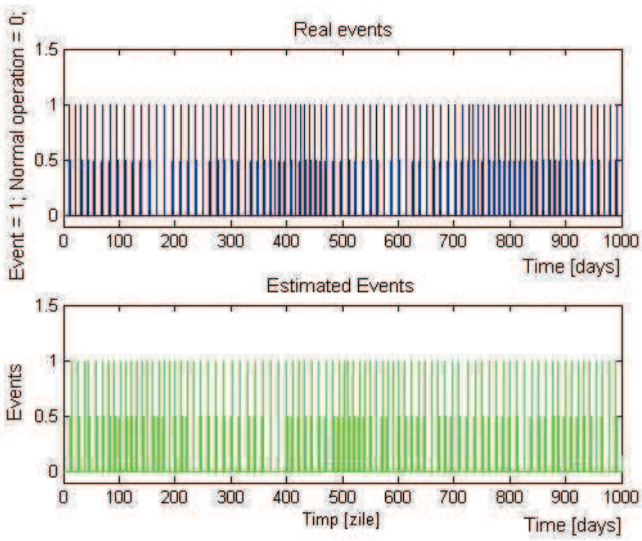

Figure 21: Event occurrence diagram for centre Oradea

5 Conclusions

Analyzing the error comparison diagrams (figures 5, 11 and 17) it can be observed that in all cases we have a convergence of values after 150 epochs. In diagram in figure 6 (centre Alesd), the predicted values are in the range of 200-400 events. The predicted values, which can be considered as satisfactory are in the range of 200-350 events.

In diagrams at figures 12 (centre Stei) and 18 (centre Oradea), the predicted values are in the interval of 300-600 events. for centre Stei, prediction values are considered satisfactory in the interval of 300-350 events and for centre Oradea, the satisfactory prediction is considered to be in the range of 300-500 events. In this range the prediction error is about 1-2 days.

The best prediction results are for centre Oradea and the worst are for centre Stei.

The event period and event occurrence diagrams presented can be considered as new tools for reliability studies in the field of electric energy systems.

The researches lead the authors to the conclusion that a good result can be obtained only if a large amount of data is available. In order to have good predictions the energy delivering centers must rethink their data recording systems to be more accurate and meaningful. The data should be also selectable in function of the event's nature. Consideration of event nature could not be made in this study because of missing information. Future studies should also consider additional information as machine and device condition and equipment age, which would improve prediction accuracy. For this approach a multi-criteria fuzzy inference system can be used.

There had been acknowledged that neuro-fuzzy prediction is a useful tool for this kind of systems reliability analysis.

References

[1] S. Dziţac, *Contribution to Modeling And Simulation of Electric Power Distribution Systems Reliability and Availability Performances*, PhD Thesis, University of Oradea, 2008.

[2] J.-S. R. Jang, "*Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm*," Proc. of the Ninth National Conf. on Artificial Intelligence (AAAI-91), pp. 762-767, July 1991.

[3] J.-S. R. Jang, "*ANFIS: Adaptive-Network-based Fuzzy Inference Systems*," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23, No. 3, pp. 665-685, May 1993.

[4] J.-S. R. Jang, and C.-T. Sun, *Neuro-fuzzy modeling and control*, Proceedings of the IEEE, March 1995.

[5] J.-S. R. Jang, and C.-T. Sun, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, 1997.

[6] C. Secui, I. Felea, S. Dziţac, E. Dale, I. Boja, *Database and software system for evaluation of quality indexes of electric energy supplying service*, Scientific Bulletin of the "PO-

LITEHNICA", Timioara, Seria Energetica, Proceedings of the 7 th International Power Systems Conference, november 22-23, PSC 2007, pag. 589-596, ISSN 1582-7194.

[7] L. A. Zadeh,, "Fuzzy Logic," Computer, Vol. 1, No. 4, pp. 83-93, 1988.

[8] *Elaborarea unui pachet software si constituirea bazei de date pentru evaluarea indicatorilor de calitate a serviciului de furnizare a EE la nivelul consumatorilor industriali si edilitari din judetul Bihor*, Raport la contractul de cercetare nr. 3050/2006, Universitatea din Oradea, Facultatea de Energetica, 2007.

Simona Dziţac
University of Oradea
Universitatii St. 1, 410087, Oradea, Romania
E-mail: sdzitac@rdslink.ro

Ioan Felea
University of Oradea
Universitatii St. 1, 410087, Oradea, Romania
E-mail: ifelea@uoradea.ro

Ioan Dziţac
Department of Economic Informatics
Agora University of Oradea
Piata Tineretului 8, Oradea 410526, Romania
E-mail: idzitac@univagora.ro

Tiberiu Vesselenyi
University of Oradea
Universitatii St. 1, 410087, Oradea, Romania
E-mail: tvesselenyi@yahoo.co.uk